

Your Justice Is Different From My Justice: Quantifying Mental State Variation of Identical Words Through a Case Study of Korean Newspaper Corpora

Hana Jee, York St John University, United Kingdom

The Asian Conference on Media, Communication & Film 2024
Official Conference Proceedings

Abstract

This study offers a quantitative investigation of a well-established question in theoretical linguistics: how identical words can carry subtle semantic variations across different contexts. While linguists have long recognized this phenomenon, empirical measurement has been elusive. Using computational linguistic techniques, this research analyzes politically contrasting two Korean newspapers as corpora to demonstrate and quantify how lexical meanings are shaped by surrounding context. Three methods were employed: Latent Semantic Analysis, Topic Modeling, and Sentiment Analysis. The findings provide empirical support for theoretical concepts of fluid word meanings. Both abstract and concrete words exhibited measurable context-dependent semantic shifts, with concrete words showing stronger sentimental biases. This approach to quantifying lexical semantics contributes to the validation of linguistic theories and opens new avenues for exploring language use across various domains. The results have potential implications for cross-cultural communication, language acquisition, and media practices.

Keywords: Lexical Semantics, Context-Dependent Meaning, Quantified Word Meaning, Korean Media

iafor

The International Academic Forum
www.iafor.org

1. Introduction

Modern linguistic theory has shifted away from viewing words as independent units stored in memory, instead conceptualizing them as activated neural networks encompassing sensory memories, experiences, and contextual knowledge (Binder & Desai, 2011; Elman, 2009). Rather than accessing fixed meanings in specific brain regions, words act as conduits to a complex cognitive web that is dynamically retrieved during processing (Taylor, 2006; Sztencel, 2018).

During language processing, the brain employs pragmatic inferences to constrain meaning, with contextual cues activating only the most relevant associations for a given situation (Grice, 1975; Barsalou, 2005). This context-dependent activation suggests that word meanings cannot be identical across individuals, but rather represent shared social agreements with room for variation (Burge, 1979; Horton & Gerrig, 2005).

This phenomenon is illustrated by Putnam's (1974) 'twin earth' thought experiment, which demonstrates how identical words can evoke different mental states based on accumulated experiences. Various theoretical frameworks have attempted to capture this semantic flexibility, including Cruse's (2004) 'facets,' Falkum's (2015) 'ad hoc concepts,' and Carston's (2019) 'polysemy complex.' In legal philosophy, Gallie's (1955) 'essentially contested concepts' similarly addresses terms that admit multiple reasonable interpretations.

This study aims to quantitatively examine these theoretical frameworks by analyzing semantic variations in politically contrasting media outlets. This approach is justified by the following: (1) different societal groups view reality through distinct lenses reflected in news media, (2) political contrasts may highlight lexical fluidity, (3) newspapers contain both abstract and concrete words, and (4) their formal nature facilitates corpus analysis.

2. Methodology

2.1. Corpus Creation and Target Word Selection

The study analyzed two Korean newspapers representing contrasting political perspectives: *Dong A* (conservative) and *Oh My News* (progressive). Data was collected via web scraping from Politics and Society sections between July and December 2022, yielding approximately 2 million words per corpus category.

Target words were selected based on three criteria: presence in both corpora, frequency above a predetermined threshold, and classification as pivotal vocabulary. The analysis included:

27 Abstract words

권력(political power), 보수(conservatism), 진보(progressivity), 권리(right), 정의(justice), 공정(fairness), 개혁(reformation), 노동(labor), 경쟁(competition), 민주주의(democracy), 복지(welfare), 사랑(love), 사상(ideology), 성장(growth), 안보(security), 연대(solidarity), 윤리(ethics), 의무(obligation), 이념(idea), 인권(human right), 자유(freedom), 책임(responsibility), 평등(equality), 평화(peace), 혁명(revolution), 혁신(innovation), 협력(cooperation)

9 Concrete words

일본(Japan), 미국(US), 영국(UK), 북한(North Korea), 대통령(president), 삼성(Samsung), 핵(nuclear), 장애인(the disabled), 친일(pro-Japanese)

For each target word, two separate corpora were created by extracting all sentences containing the word from each news source. This enabled analysis of how each outlet constructs meaning through contextual usage. The analysis employed three text-mining approaches: LDA Topic Modelling, Latent Semantic Analysis, and Sentiment Analysis. The first two methods identified associated terms, while sentiment analysis determined positive/negative portrayals. Figure 1 illustrates the overall process.

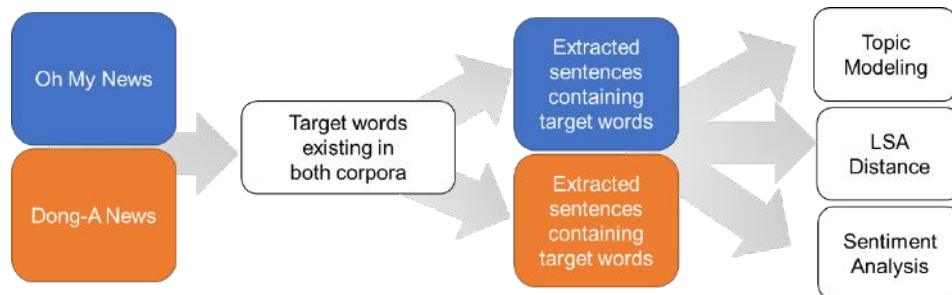


Figure 1: Visual Overview of the Methodology Process

2.2. Analysis Techniques

2.2.1. Latent Semantic Analysis (LSA)

LSA extracts word meaning by analyzing contextual environments (Landauer, Foltz, & Laham, 1998). Using singular value decomposition (SVD), it reduces dimensionality to simulate human neural networks, effectively detecting latent relationships between words even when they don't directly co-occur. This technique has been validated through various behavioral research, including information retrieval and semantic priming studies. I used genism Word2Vec model (parameters: size=100, window=5, min_count=1, workers=4) to extract 100 semantically closest words for each target word. Visualization was performed using networkX in Python 3.7. The full script is available from https://github.com/HanaJee/LSA_semantic-distance_Korean-corpus.

2.2.2. Latent Dirichlet Allocation (LDA) Topic Modelling

LDA Topic Modelling (Blei, Ng, & Jordan, 2003) identifies multiple topics within documents, with each topic comprising words of varying probabilities. While traditionally applied to coherent narratives, this study uniquely used it to analyze sentence collections containing target words, focusing on identifying repetitive media frames. I applied genism Ldamodel (parameters: num_topics=5) with CoherenceModel for topic coherence evaluation. The full script is available from <https://github.com/HanaJee/LDA-topic-modeling/blob/main/code>.

2.2.3. Sentiment Analysis

Sentiment analysis identifies emotional content in text (Liu, 2012) through two main approaches. The lexicon-based method uses established emotion dictionaries (Ding, Liu, &

Yu, 2008), but shows limited accuracy for Korean (Park, 2017) due to the language’s agglutinative nature and high frequency of homonyms (Kang, 2005).

The alternative machine learning approach uses algorithms trained on extensive corpora (Mullen & Collier, 2004), incorporating deep learning techniques (Seo & Kim, 2016). This study employed Tensorflow LSTM (parameters: for compiling, optimizer=‘rmsprop’, loss=‘binary_crossentropy’, metrics=[‘acc’]; for fitting, epochs=15, batch_size=60, validation_split=0.2), trained on Naver movie reviews with 15 epochs (batch_size=60, validation_split=0.2), achieving over 90% accuracy. The full script is available from <https://github.com/HanaJee/Sentiment-Analysis>.

3. Results

This study investigated whether semantic variation of identical words can be quantified by context and if patterns differ between abstract and concrete words. The analysis revealed quantifiable semantic variations, supporting theories of context-sensitive word meanings (Cruse, 2004; Falkum, 2015; Carston, 2019). Using politically contrasting newspapers highlighted distinct media frames (Goffman, 1974; Gitlin, 2003).

Topic modelling revealed that conservative Dong A frequently emphasizes national security, while progressive Oh My News focuses more on societal communication, reflecting their political leanings (Choi, Hong, & Jho, 2021). The broader topic range in Oh My News likely stems from its citizen journalism format. Figure 2 and 3 exemplify two salient words, ‘peace’ as an abstract word and ‘the U.S.’ as a concrete word, which show clear semantic contrast between two media outlets. These show how identical words emphasize certain facets depending on contexts. Table 1 below includes 11 abstract words and 3 concrete words demonstrating clear difference in nuance of identical words.

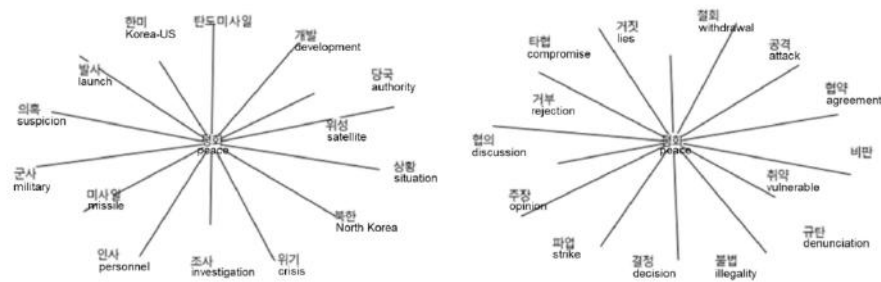


Figure 2: Conservative (Left) and Progressive Definition of ‘Peace’ (Right)

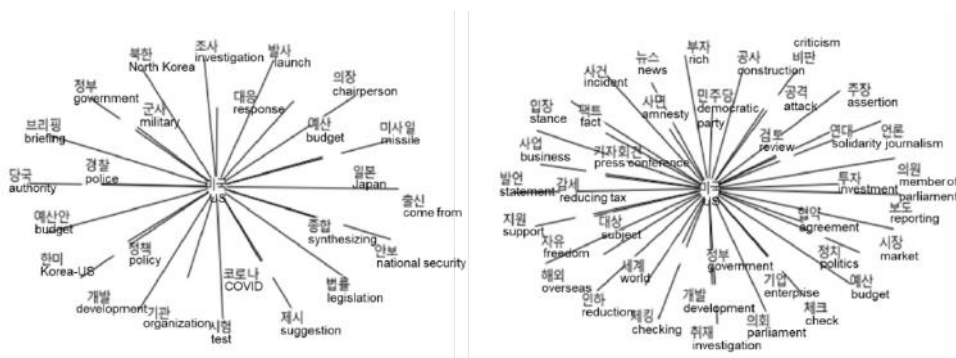


Figure 3: Conservative (Left) and Progressive Definition of ‘U.S.’ (Right)

Table 1. Re-defined Words Based on Each Media Outlet

Abstract words	Conservative (Dong A)	Progressive (Oh My News)
개혁 Reformation	Associated with strengthening national security and government response to threats, by means of increased military spending, US cooperation, and bolstering the police force . Driven by the ruling party and the Ministry of the Interior and Safety, these initiatives are shaped by National Assembly legislation and budgets .	Associated with promoting social justice , civil liberties , and democratic accountability , by means of press conferences, media reports, and fact-finding missions that criticize oppressive government policies. The reforms advocate for cooperation between authorities and measures like pardons for the unjustly imprisoned, despite conservative resistance.
권력 Political power	To maintain national security, particularly regarding the US alliance and North Korea's threats, efforts focus on strengthening law enforcement agencies and revising laws to enhance government protection . This includes empowering the police and the Ministry of the Interior and Safety to respond effectively to security incidents.	Associated with corporate and wealthy influence undermining democratic principles and fair economic policies, issues include business tax rates , corporate regulations, and unethical practices . The role of lawyers, media, and elections is emphasized ensuring accountability and promoting transparency and democratic values.
노동 Labour	Associated with government budget allocation, law enforcement, and national security, the focus is on the ruling party, government ministers, and police managing labour-related matters to maintain stability and order . The President's office and high-ranking police officials play key roles in shaping labour policies and responding to potential disruptions or emergencies .	Associated with economic inequality , corporate power, and workers' rights , the focus is on how cities, media, and government budget and tax policies impact labour conditions and worker welfare . Business tax cuts, labour movement suppression, and corporate influence on labour practices are criticised by press conferences and fact-finding efforts to expose injustices and advocate for workers' rights to strike and protest .
민주주의 Democracy	Associated with government institutions and elected officials representing the people's will, the focus is on government responses and political engagement . The National Assembly, government committees, and the ruling party shape democratic outcomes through discussions and decision-making.	Associated with protecting civil liberties , press freedom , and holding power accountable, the focus is on media truthfulness , political processes like primaries, budgets, and tax policies in promoting democracy. Citizen solidarity , fact-checking , and scrutinizing government claims are emphasized.
성장 Growth, development	Primarily associated with strengthening national security and the US military alliance to counter North Korean threats, the focus is on government, law enforcement, and military leaders' actions in protecting the nation. A strong security environment is seen as crucial for growth, upheld through legislation and state power .	Associated with economic development , social welfare , and political empowerment, the focus is on supporting businesses, investments , and regional projects, and allocating budgets for welfare programs and infrastructure . Growth ties to election outcomes, party strength, and media influence. Media freedom and government support for urban development is stressed to foster economic and social progress.

<p>안보 National security</p>	<p>Primarily focused on military threats from North Korea, particularly its missile activities, the emphasis is on the government's responsibility to gather intelligence and take decisive action with the US and Japan. The role of law enforcement agencies, like the police, is also crucial in maintaining national security.</p>	<p>Security as economic and social, emphasizing stable incomes and living standards in a connected world. Linked to income security, global cooperation, and government responses to challenges, the focus is on State Council discussions and resource mobilization for security beyond military threats.</p>
<p>인권 Human right</p>	<p>Discussed in parliamentary investigations, the focus is on government officials, political dynamics, and human rights, often framed in terms of national security and crisis management. The actions of National Assembly figures, ministers, and the President's office, as well as police promotions and investigations, are key aspects of this discourse.</p>	<p>Associated with protecting vulnerable groups and promoting social and economic justice, advocacy for individual rights and addressing economic inequalities is emphasized. Media reporting on human rights issues is stressed, with accountability demanded from those in power, including calls for the release of prisoners of conscience and unjustly convicted individuals, extending concerns to other countries like Vietnam.</p>
<p>자유 Freedom</p>	<p>Associated with economic policies, foreign relations, and national security, the focus is on fostering economic growth and security through strategic investments and diplomacy. Strong leadership, national interests and values like human rights are prioritized, addressing challenges such as low birth rates and military expansion.</p>	<p>Closely tied to the media and civil society, progressives advocate for press freedom to ensure transparency and accountability. Fact-checking and freedom of expression are emphasized as vital for democracy, prioritizing issues like pardons for political prisoners and measures to address inequalities.</p>
<p>책임 Responsibility</p>	<p>Associated with government responsibility in handling issues, the focus is on investigating and addressing problems with accountability and improvement. Thorough inquiries are emphasized to uncover truth and hold the responsible accountable, expecting proactive government action and transparent communication with the public.</p>	<p>Closely tied to the media and civil society, accountability is emphasized through critical analysis and truth-telling. The media exposes wrongdoing and demands change, highlighting injustices and advocating solutions. Authorities are urged to take responsibility and address public concerns.</p>
<p>평화 Peace</p>	<p>Efforts to reduce tensions and prevent escalation through diplomacy and negotiations are paramount. Amid military tensions and weapons development, concerns and suspicions are raised by North Korea's ballistic missile and satellite activities, posing significant threats to regional peace.</p>	<p>Efforts to resolve conflicts in contentious situations, such as labour disputes or political controversies involving criticism, false claims, and intense actions, are undertaken through dialogue and compromise. Despite challenges and the presence of inflammatory actions, these efforts aim to find common ground and promote resolution.</p>
<p>협력 Cooperation</p>	<p>Associated with government efforts to manage challenges through collaboration, focusing on policies for security, development, and international relations. Informed decision-making and cooperation are emphasized to tackle complex issues, expecting collaboration with stakeholders to promote national interests.</p>	<p>Solidarity is emphasized in tackling social, economic, and political issues like inequality and environmental degradation, particularly across government levels and sectors, supported by media accountability. Crucial for positive change and promoting societal well-being.</p>

Concrete words	Conservative (Dong A)	Progressive (Oh My News)
미국 The U.S.	Discussed primarily in terms of national security and military cooperation , the focus is on its role in addressing North Korean threats , emphasizing the importance of the US-Korea alliance for regional stability and defence capabilities. This extends to collaboration on law enforcement, intelligence sharing, and global challenges like the COVID-19 pandemic .	Discussed in economic and international contexts , the focus is on the impact of its investments and policies, with progressives advocating for accountability and social justice . The role of the media in accountability is emphasized, government actions are criticized, and international cooperation on global challenges like climate change and human rights abuses is prioritized.
북한 North Korea	Primarily discussed as a military threat , the focus is on its missile launches and weapons development , prompting a robust response from South Korea and its allies. Intelligence gathering and cooperation with the United States and Japan is stressed, expecting decisive action from the government, including resource allocation and support for defence efforts.	Calls for a more engaging approach to inter-Korean relations often involve criticism of the South Korean government. Emphasizing dialogue and economic cooperation to build trust, concerns about human rights in North Korea are also raised, alongside criticism of government budget priorities and tax policies, with questions about appointments and legislation handled by the media .
일본 Japan	Mentioned in relation to security issues and cooperation with South Korea and the United States, its role in addressing the North Korean threat is emphasized, particularly in intelligence sharing and coordinated responses. Its involvement in domestic politics is also discussed, including policies and actions of the ruling party and government officials, alongside budget allocation and crisis management.	Critiques of its conservative policies focus on concerns about budget priorities and media transparency. Advocacy for marginalized communities and civil liberties is emphasized, with the media playing a vital role in informing the public and advancing causes.

Sentiment analysis revealed distinct patterns between abstract and concrete words. Abstract words showed balanced emotional usage across both outlets (Figure 4), while concrete words demonstrated stronger sentimental bias (Figure 5). Notable examples include conservative Dong A’s negative portrayal of ‘North Korea’ (62% negative) and Oh My News’ negative representation of ‘Japan’ (61% negative), both with over 90% algorithmic confidence.

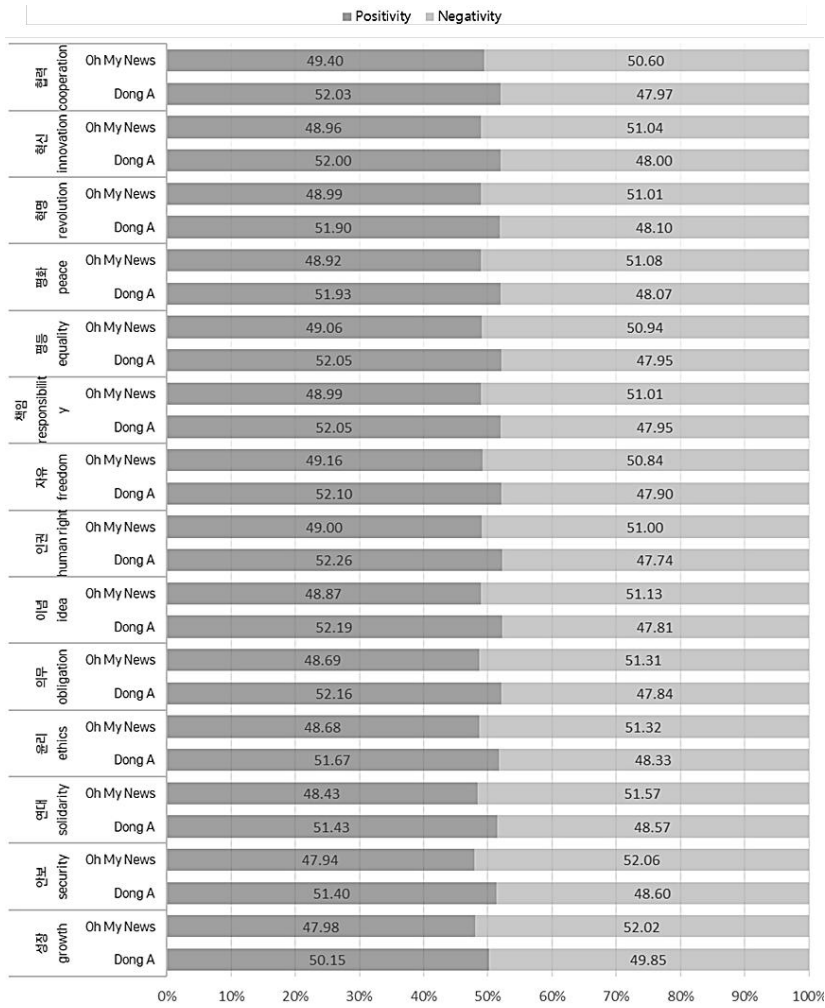


Figure 4: The Percentage (%) of Sentences With Positive and Negative Sentiment for Each Target Abstract Word

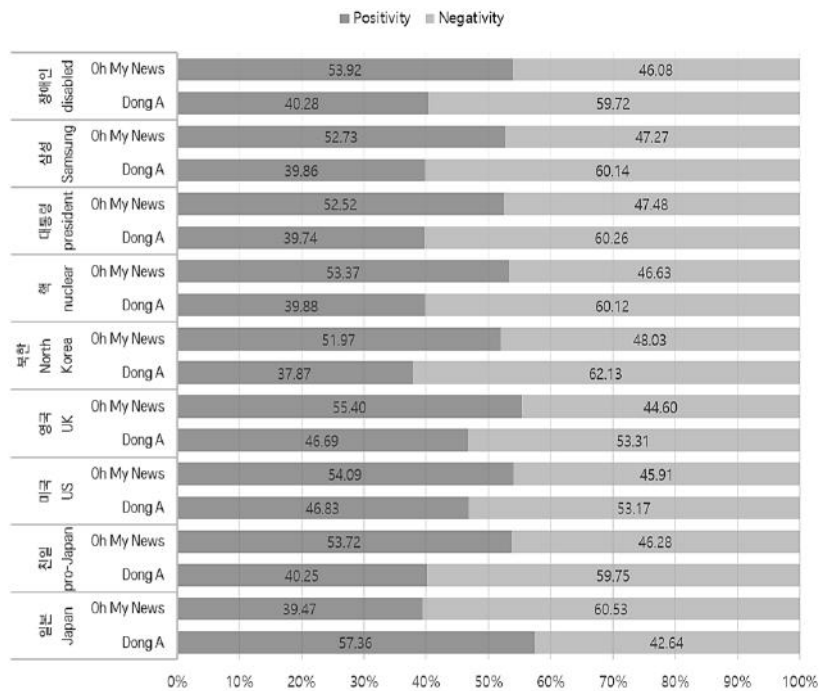


Figure 5: The Percentage (%) of Sentences With Positive and Negative Sentiment for Each Target Concrete Word

In conclusion, each method provided unique insights: LSA revealed context-dependent meaning emphasis, topic modelling confirmed distinct media frames, and sentiment analysis showed different emotional handling of abstract versus concrete words.

4. Discussion

As an important proof of concept, this study is the first of its kind seeking to empirically measure subtle shifts in lexical semantics using computational techniques. Based on the two politically contrasting newspaper outlets, it was confirmed that certain sides of lexical semantics can be emphasized by its contexts around it. A significant number of the target words demonstrated quantifiable differences in meaning between the conservative and progressive newspaper corpora. The fact that the progressive newspaper includes citizen journalists may have influenced the width and depth of topics covered.

The findings carry significance for both theoretical linguistics and practical applications. On the theoretical side, it provides the first quantitative validation of concepts like ‘facet’ (Cruse, 2004), ‘ad hoc concept’ (Falkum, 2015), or ‘polysemy complex’ (Carston, 2019) that capture variability in word meaning, lending credence to proposals about the fluid, context-sensitive nature of semantics. The computational linguistic methods establish an empirical approach to reveal and analyse meaning nuances, enabling expansion to other phrases and constructions.

Unlike abstract words, concrete words exhibit a stronger reliance on the external world and can therefore be experienced without the mediation of language. They are influenced or compromised by other concepts to a lesser degree (Weimer-Hastings & Xu, 2005). The current study showed that both abstract and concrete words can be shaped by their contexts, possibly due to the same method (i.e. parameters) was applied to them. It lends further quantitative validation to the lack of rigid lexical concepts, now expanded across word types.

Practically, the findings imply the possibility of communicating with identical vocabulary that carries different meanings, especially within cross-cultural contexts. As debates intensify amid democratization, global policy issues require shared understanding between disparate worldviews mediated by language. This research prompts more responsible, unifying communication by heightening awareness of how subtly words can polarize and impede vital consensus.

When it comes to learning, the findings imply the profound impact of frequent linguistic context on constructing lexical meanings early in life. The human brain possesses a remarkable capacity for plasticity, allowing it to adapt and change in response to experiences and environmental stimuli (Abbott & Nelson, 2000; Feldman, 2008). When neurons are frequently activated by specific contexts, the synaptic connections between them are strengthened, facilitating the retrieval and processing of associated concepts and ideas (Dabrowska, 2004; Hebb, 1949). Consequently, repeated exposure to biased contexts can profoundly influence the development of lexical meanings, particularly in young individuals whose cognitive structures are still being established (Barsalou, 2005; Gibbs, 2002).

As these biased associations become more robust and easily accessible, they may overshadow alternative semantic possibilities, potentially limiting the individual’s ability to tap into a broader range of semantic resources. This underscores the importance of promoting diverse and balanced linguistic environments to foster the development of flexible and adaptable lexical networks, especially during the formative years of language acquisition.

5. Conclusion

This study provides the first empirical evidence for measuring subtle semantic variations in word meaning through computational techniques. The findings validate theoretical concepts of semantic flexibility while demonstrating that both abstract and concrete words can be shaped by their contexts. Beyond theoretical implications, this research highlights the critical role of linguistic context in meaning construction, particularly in cross-cultural communication and early language acquisition. This underscores the importance of fostering diverse linguistic environments, especially during formative years, to promote flexible semantic understanding and facilitate more effective cross-cultural dialogue.

Ongoing research directions offer several promising avenues: exploring evidence and strategies on how media manipulate ideas and public attitudes through language use; cross-cultural differences in lexical meanings; and a broader range of qualitative methodologies. Enriched by more robust empirical substantiation, these efforts could collectively foster heightened public awareness, encouraging a more conscientious approach to language use in both perception and expression.

References

- Abbott, L. F., & Nelson, S. B. (2000). Synaptic plasticity: taming the beast. *Nature neuroscience*, 3(11), 1178-1183.
- Barsalou, L. W. (2005). Continuity of the conceptual system across species. *Trends in Cognitive Sciences*, 9(7), 309-311.
- Binder, J. R., & Desai, R. H. (2011). The neurobiology of semantic memory. *Trends in cognitive sciences*, 15(11), 527-536.
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of machine Learning research*, 3(Jan), 993-1022.
- Burge, T. (1979). Individualism and the Mental. *Midwest studies in philosophy*, 4, 73-121.
- Carston, R. (2019). *Ad hoc concepts, polysemy and the lexicon*. Cambridge University Press.
- Choi, J. D., Hong, S. M., & Jho, W. S. (2021). Framing Contests over International News: An Analysis of Media Discourse on GSOMIA Termination. *The Korean Association For Information Society*, 22(2), 51-75.
- Cruse, D. A. (2004). Lexical facets and metonymy. *Ilha Do Desterro a Journal of English Language, Literatures in English And Cultural Studies*, (47), 073-096.
- Dabrowska, E. (2004). *Language, mind and brain: Some psychological and neurological constraints on theories of grammar*. Edinburgh University Press.
- Ding, X., Liu, B., & Yu, P. S. (2008, February). A holistic lexicon-based approach to opinion mining.
- Elman, J. L. (2009). On the meaning of words and dinosaur bones: Lexical knowledge without a lexicon. *Cognitive science*, 33(4), 547-582.
- Falkum, I. L. (2015). The how and why of polysemy: A pragmatic account. *Lingua*, 157, 83-99.
- Feldman, J. (2008). *From molecule to metaphor: A neural theory of language*. MIT press.
- Gallie, W. B. (1955, January). Essentially contested concepts. In Proceedings of the Aristotelian society (Vol. 56, pp. 167-198). *Aristotelian Society*, Wiley.
- Gibbs Jr, R. W. (2002). A new look at literal meaning in understanding what is said and implicated. *Journal of pragmatics*, 34(4), 457-486.
- Gitlin, T. (2003). *The whole world is watching: Mass media in the making and unmaking of the new left*. Univ of California Press.
- Goffman, E. (1974). *Frame analysis: An essay on the organization of experience*. Harvard University Press.

- Grice, H. P. (1975). Logic and Conversation. *Syntax and Semantics*, 3, 43-58.
- Hebb, D.O. (1949). *The Organization of Behavior*. New York: Wiley & Sons.
- Horton, W. S., & Gerrig, R. J. (2005). The impact of memory demands on audience design during language production. *Cognition*, 96(2), 127-142.
- Kang, B. M. (2005). Aspects of the use of homonyms. *Language research* 41 (1), 1-29.
- Landauer, T. K., Foltz, P. W., & Laham, D. (1998). An introduction to latent semantic analysis. *Discourse processes*, 25(2-3), 259-284.
- Liu, B. (2012). Sentiment analysis and opinion mining. *Synthesis lectures on human language technologies*, 5(1), 1-167.
- Mullen, T., & Collier, N. (2004). Sentiment analysis using support vector machines with diverse information sources.
- Park, S. (2017). Sentimental Analysis of SW Education News Data. *Journal of The Korean Association of Information Education*, 21(1), 89-96.
- Putnam, H. (1974). Meaning and reference. *The journal of philosophy*, 70(19), 699-711.
- Seo, S. H., & Kim, J. T. (2016). Research pattern of deep-learning-based sentiment analysis.
- Sztencel, M. (2018). Semantics, Pragmatics and Meaning Revisited.
- Taylor, K. (2006). *Brainwashing: The science of thought control*. Oxford University Press.
- Weimer-Hastings, K., & Xu, X. (2005). Content differences for abstract and concrete concepts. *Cognitive science*, 29(5), 719-736.